

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:02 ; Search time 210.42 Seconds
(without alignments)
108.227 Million cell updates/sec

Title: US-09-331-631A-3
Perfect score: 3532
Sequence: 1 MAINTSNLCSLFLSLFL.....SSRSRKQOOPVSLDPVGF 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_36.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3532	100.0	666	19 W62829	Macadamia integrifolia
2	3412	96.6	666	19 W62828	Macadamia integrifolia
3	3215	91.0	625	19 W62830	Macadamia integrifolia
4	1145	32.4	566	13 R20181	Sequence encoded b
5	1105.5	31.3	590	19 W62832	Cosyptium hirsutum
6	1036	29.3	525	19 W62831	Theobroma cacao an
7	895	25.3	593	19 W62835	Zea mays antimicro
8	875	24.8	637	19 W62837	Hordeum vulgare an
9	873.5	24.7	614	18 W2149	Peanut allergen Ar
10	873.5	24.7	614	18 W62834	Archis hypogaea a
11	865.5	24.5	626	20 Y15244	Peanut allergen A
12	865.5	24.5	626	20 Y25657	Peanut allergen I1

13	859.5	24.3	626	18 W2150	Peanut allergen Ar
14	855.5	24.2	605	19 W62838	Glycine max anti
15	839.5	23.8	605	20 Y40999	Soybean beta-congl
16	798	22.6	524	20 W90339	G. max SBP1 protei
17	797.5	22.6	489	20 W90341	G. max SBP2 protei
18	739	20.9	444	20 W90340	G. max truncated S
19	727.5	20.6	415	20 Y40913	Ara h 1 allergen p
20	689.5	19.5	409	20 W90342	G. max truncated S
21	530.5	15.0	371	20 Y40914	T. gondii immunoge
22	237	6.7	611	20 Y29039	A human trichothal
23	226	6.4	1898	20 Y30795	Human thyrotropin
24	218.5	6.2	1162	21 Y58500	Human follicle sti
25	187.5	5.3	562	16 R70491	Modified oat globi
26	184	5.2	412	17 W03626	Rice storage prote
27	183.5	5.2	1239	20 Y55931	Peanut allergen, A
28	183	5.2	360	17 W03627	Ara h 3 allergen s
29	173	4.9	472	15 R47127	P. falciparum LSA-R
30	171	4.8	499	9 P82755	Mouse STE20-relate
31	170	4.8	510	20 Y15246	Human CDC28-#3 RNA
32	169	4.8	512	20 Y40912	Protein regulat
33	166	4.7	493	13 R26944	Oat globin A2B. A
34	164.5	4.7	316	13 R26941	Human ZC3 protein.
35	163.5	4.6	1233	20 Y55954	Human GEX2 protein
36	163	4.6	1135	21 Y68784	Biorhyth marker p
37	161	4.6	1041	20 W36613	Human ZC2 protein.
38	160	4.5	1299	21 Y58633	LexA/NIMA fusion p
39	159.5	4.5	515	15 R47128	
40	157	4.4	968	20 Y55966	
41	154.5	4.4	1326	20 Y55933	
42	147	4.2	912	20 Y55939	
43	146.5	4.1	361	20 Y22206	
44	146.5	4.1	1297	20 Y55932	
45	146.5	4.1	2192	18 W21732	

ALIGNMENTS

RESULT	1
ID	W62829 standard; Protein: 666 AA.
XX	
AC	W62829;
XX	
DT	27-OCT-1998 (first entry)
XX	
DE	Macadamia integrifolia antimicrobial protein.
XX	
KW	antimicrobial protein; infestation; control.
XX	
OS	Macadamia integrifolia.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	/note="signal peptide"
FT	29..666
FT	/note="mature protein"
XX	
PN	MO9827805-A1.
XX	
PD	02-JUL-1998.
XX	
PF	22-DEC-1997; 97WO-AU00874.
XX	
PR	20-DEC-1996; 96AU-0004275.
XX	
PA	(RETR-) COOP RBS CENT TROPICAL PLANT PATHOLOGY.
XX	
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP.
XX	
DR	WPI: 1998-377279/32.
XX	
DR	N-PSDB: V42311.
XX	

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX useful for controlling microbial infestations of plants or mammals
XX
XX Claim 1: Page 39-41; 96pp; English.
XX
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX
XX Sequence 666 AA:
SO

Query Match 100.0%; Score 3532; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAINTSNLCSLFLSLFLSTTVSLAESFDRQYEECKRQCMOLETSQMRVCSQCD 60
DB 1 maintsnlcsllflslstvtvlaesefdrqyeecrkqcmqltsqmrvcvsqd 60
QY 61 KRFEEDIDMSKYNDODPQTDCCQCRRCROESGPROOQYQCRCKEICEEEYNNQR 120
DB 61 krfeedidwskyndqddpqtdccqcrtrcqsgsprqqycqrckelceeeeynnqr 120
QY 121 DPQOQTECCQCRQRETEPRHMOTCOQRCERYEKKRKQKRYEQQREDEKYEEM 180
DB 121 dpqdyeqcqrctqrheteprhmotcqrcteryekrkqkryeqqredeekyeem 180
QY 181 KEEDNKRDPQREYEDCRRCEQOEPRQOYQCRCEQORHGRGDDLPPQSGSGRY 240
DB 181 keednkrdpqreyedcrrceqoeprqyqqrceqgrqyrgddlmpqsgsgr 240
QY 241 EECEKQSDNPYFFDEKSLSTRTEEGHISVLENFYGSKLRLAKNRYLLEANPNA 300
DB 241 eegekqsdnpyffderslsttrteeeghlsvlenfygskllralknyllleapna 300
QY 301 FVLPTHLDADAILLVYGRGALKMIRHNRRESYNLECGDVIRIPAGTTFYLIRDNNEKL 360
DB 301 fvlpthldadaillvgrgalkmihndresynlecgdviripagtlfyllrndnertl 360
QY 361 HIAKFLQITSTPGQYKEFPFAGGONPEPYLSTFSKELALMTQAEKRLGVAGQOOREGV 420
DB 361 hiakfqltistpgykeffpagggnpepylstskelleaaltgaerlrgvlgqoregv 420
QY 421 IISASQEQIRELTRDSESRHWHIRRGESSRGPYLVFNKRPLYSNKYGOAYEVKPEDYR 480
DB 421 iisaseqireltrdsestrwhirrgessrgpylvfnkrplysnkygoyevkpedyr 480
QY 481 QLODMYVSFIANITOGSSMMGFNTBSTRKVVVYVAGSADYENACPHLSGRHGRGGR 540
DB 481 qlqdmvsvfianltqgsmngpfntbstrkvvvvasgadvenacphlsgrhgrggr 540
QY 541 HEEEDVHYEQVAKRLSKREAIIVPVGHPVVFSSGNEMLLFAPGINAQNHNENFLAGR 600
DB 541 heeedvhyeqvakrlskrealivpvghpvvfssgneillfapginaqnhenflagr 600
QY 601 ERNVLOQIEPOAMELFAAPRKEVEELFNSODESIFFPGRHOHQOOSRSRSTKQOOLVST 660
DB 601 ervnlqiepqamelafaprkeveelfnsodesiffpgrhohqoosrsrstkqolvs 660
QY 661 LDFVGF 666
DB 661 ldfvgf 666

RESULT 2
ID W62828 standard; Protein: 666 AA.
XX
AC W62828;
XX
DT 27-OCT-1998 (first entry)
XX

DE Macadamia integrifolia antimicrobial protein.
XX
XX antimicrobial protein; infestation; control.
XX
XX Macadamia integrifolia.
XX
XX
XX
XX
XX Key Location/Qualifiers
FT Peptide 1..28
FT Protein /note= "signal peptide"
FT Protein 29..666
FT Protein /note= "mature protein"
XX
XX W09827805-A1.
XX
XX 02-JUL-1998.
XX
XX 22-DEC-1997; 97WO-AU00874.
XX
XX 20-DEC-1996; 96AU-0004275.
XX
XX (HETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP.
XX
XX WPI; 1998-377279/32.
XX
XX DR N-PSDB; V42310.
XX
XX
XX
XX Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
XX Claim 1: Page 34-36; 96pp; English.
XX
XX
XX The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX
XX Sequence 666 AA:
SO

Query Match 96.6%; Score 3412; DB 19; Length 666;
Best Local Similarity 96.1%; Pred. No. 1.5e-298;
Matches 640; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 MAINTSNLCSLFLSLFLSTTVSLAESFDRQYEECKRQCMOLETSQMRVCSQCD 60
DB 1 maintsnlcsllflslstvtvlaesefdrqyeecrkqcmqltsqmrvcvsqd 60
QY 61 KRFEEDIDMSKYNDODPQTDCCQCRRCROESGPROOQYQCRCKEICEEEYNNQR 120
DB 61 krfeedidwskyndqddpqtdccqcrtrcqsgsprqqycqrckelceeeeynnqr 120
QY 121 DPQOQTECCQCRQRETEPRHMOTCOQRCERYEKKRKQKRYEQQREDEKYEEM 180
DB 121 dpqdyeqcqrctqrheteprhmotcqrcteryekrkqkryeqqredeekyeem 180
QY 181 KEEDNKRDPQREYEDCRRCEQOEPRQOYQCRCEQORHGRGDDLPPQSGSGRY 240
DB 181 keednkrdpqreyedcrrceqoeprqyqqrceqgrqyrgddlmpqsgsgr 240
QY 241 EECEKQSDNPYFFDEKSLSTRTEEGHISVLENFYGSKLRLAKNRYLLEANPNA 300
DB 241 eegekqsdnpyffderslsttrteeeghlsvlenfygskllralknyllleapna 300
QY 301 FVLPTHLDADAILLVYGRGALKMIRHNRRESYNLECGDVIRIPAGTTFYLIRDNNEKL 360
DB 301 fvlpthldadaillvgrgalkmihndresynlecgdviripagtlfyllrndnertl 360
QY 361 HIAKFLQITSTPGQYKEFPFAGGONPEPYLSTFSKELALMTQAEKRLGVAGQOOREGV 420
DB 361 hiakfqltistpgykeffpagggnpepylstskelleaaltgaerlrgvlgqoregv 420
QY 421 IISASQEQIRELTRDSESRHWHIRRGESSRGPYLVFNKRPLYSNKYGOAYEVKPEDYR 480
DB 421 iisaseqireltrdsestrwhirrgessrgpylvfnkrplysnkygoyevkpedyr 480

```
Db 421 lrrsqegirelrrdsserhwhlrrgessrpylnlfnkrlpysnkyygagvewkpedyr 480
QY 481 QLODMADVVFANITOGSMGPFNTRSTKVVVVASGEADVEMACPHLSGRHGRGGR 540
Db 481 qlqmdlsvfianltqgsmgpfnttrstkvrvvasgeadvemacphlsgrhgrggr 540
QY 541 HEEEDVHYEOKARLSKREAIIVPVGHVYVSSGNENLLLEAFGINAQNHNENLAGR 600
Db 541 heeedvhyeqvarliskrealivlaghpvfvassgneilllfafiginaqnhenflagr 600
QY 601 ERNVLQOIEPQAMELAAAPRKEVEELFNSODESIFPPGPHOHOQSSRSRKQOQPLVSI 660
Db 601 ernvlqiepqamelaataaprkeveesfnsqdsilffpgrqbhgqspstlkqqplvsi 660
QY 661 LDFVGF 666
Db 661 ldfvgf 666

RESULT 3
W62830
ID W62830 standard; Protein; 625 AA.
XX
AC W62830;
XX
DT 27-OCT-1998 (first entry)
XX
DE Macadamia integrifolia antimicrobial protein.
XX
KM antimicrobial protein; infestation; control.
XX
OS Macadamia integrifolia.
XX
FH Key Location/Qualifiers
FT Peptide 1..28
FT Protein /note="signal peptide"
FT 29..666
FT /note="mature protein"
XX
PN W09827805-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-AU00874.
XX
PR 20-DEC-1996; 96AU-0004275.
XX
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR MPI; 1998-377279/32.
XX
DR N-PSDB; V42316.
XX
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX
PS useful for controlling microbial infestations of plants or mammals
XX
PS Claim 1; Page 43-45; 96pp; English.
XX
CC The sequence is that of an antimicrobial protein which can
XX
CC be used to control microbial infestations in plants and mammalian
XX
CC animals.
XX
SQ Sequence 625 AA;

Query Match 91.0%; Score 3215; DB 19; Length 625;
Best Local Similarity 96.6%; Pred. No. 7.4e-281;
Matches 604; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 42 QCMOLETSGGMRGVSGCGRPFEDIDMSKYDNODDPTDCOCORRCRQESGPRQOY 101
Db 1 qcmaletsggmrvscgckrfeeldlowskydnbdpqtecqgcrcrqesdprqy 60
```

```
QY 102 CORCKEICEEEYNNRDRPOOYEDQCEQCRQHETEPHMQTCQRCRRYERKERKQ 161
Db 61 cqrckceiceeeeynnrdrpqyqeqcqrkrctecrphmqtcqrcrrykerkrkq 120
QY 162 QKRIEEOQREDEEKYERMAKEDNKRDPQOREYEDCRRCEQOEPROOYCCORRCRQOR 221
Db 121 qkryeeqredeekeeyermakgdnkrdpqgreycdrcheqgeprtygcqrcrqeqqr 180
QY 222 OHGRCGLINPORGSGRYEEGEKOSDNVYDERLSRFRTEEGHISVLNPNFRGSK 281
Db 181 qhgrrgqlmnpqrgysryeegekgsdnpylderslsrfrteeghsvlanfygrsk 240
QY 282 LLRALKNYRLVLEANNFAVLPTHLDDAILLVTGGRGALKMIHRDNRESYNLECGDVI 341
Db 241 llralknryrlvleannfaavlpthlddadillvlggrgalkmihdntresynlecgdvi 300
QY 342 RIPAGTTFYILNRDNNEHLIAKFLQTIPTPGYKEFFPAGGONPEPYLSTFSKEILEAA 401
Db 301 ripagltfyllnrddnerlhiakflqlstipgykeffpaggonpepylstfskeillea 360
QY 402 LNNQAEFLRGVILGOORGVITISASOEQIRELTRDSSRRWHIRGESSRGPYNLFNKR 461
Db 361 lnnqaeflrgvlgoorgvitissaeqireltrdssrrwhlrrgessrgpylnlfnkr 420
QY 462 PLYSNKYGAVEYKPEDYRLOQMDVSVFANITOGSMGPFNTRSTKVVVVASGEADV 521
Db 421 plysnkygayerkpedyrlqmdvsvfianltqgsmgpfnttrstkvrvvasgeadv 480
QY 522 EMACPHLSGRHGRGGRKHEEEDVHYEOKARLSKREAIIVPVGHVYVSSGNENLL 581
Db 481 emacphlsgrhgrggrkheeeevhyeqvarliskrealivlaghpvfvassgneill 540
QY 582 LPAFGINAQNHNENFLAGRENNVLOOIEPQAMELAAAPRKEVEELFNSODESIFPPGR 641
Db 541 lfafiginaqnhenflagrerlvlqiepqamelaataaprkeveelnsqdsilffpgr 600
QY 642 OHQOQSSKSTKQOQPLVSIIDFVGF 666
Db 601 qhgqgsprstlkqqplvsiidfygf 625

RESULT 4
R20181
ID R20181 standard; Protein; 566 AA.
XX
AC R20181;
XX
DT 16-APR-1992 (first entry)
XX
DE Sequence encoded by 67 kD T. cacao protein cDNA.
XX
KM Cocoa; flavour; vicillin; seed storage protein.
XX
OS Theobroma cacao.
XX
PN W09119801-A.
XX
PD 26-DEC-1991.
XX
PF 07-JUN-1991; 91WO-GB00914.
XX
PR 11-JUN-1990; 90GB-0013016.
XX
PA (MRSC ) MARS UK LTD.
XX
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
XX
DR MPI; 1992-024418/03.
XX
DR N-PSDB; Q20377.
XX
PT Recombinant cocoa proteins - are responsible for flavour in cocoa
beans and produced in large quantities using yeast and bacterial
```


[illegible]

Db	126	georgy	-----hghnsgse-----	reergrtgrlrghegeceee	grtg	165
Qy	239	RYEEGE--EKOSDN--	PRYDEEDSLSTRFRT	EEGHISVLNENFGRSK	LBALNRYRLVL	294
Dd	166	rtgeerdeegdsrtpry	fgfrtfrllgsdnh	fvralprfdqvsrl	llgrldyvalm	225
Qy	295	EANPAFVLPRIHADAIL	LWTGSRGALKMTIHR	ONRESYNLECCDVI	RIPTGTFVLI	354
Dd	226	enprativpgrltdad	gvgyvagevyltv	lengektsytkveg	dvltvapagslmb	285
Qy	355	DNNEELHTAKFLQTI	TPRGQYKEFPFPA	GCONPERVLTSTF	SKELLFALINTQ	AERLRGVLG
Dd	286	dgrtfltiakllhtis	ypgkf-qfl-----	svkpllaslskv	lraafksdsder	letlfn
Qy	415	QO-----REGVIIS	ASQEQRLRELT	RDDSE---SRMH	IRRGCESSKCP	NLTKRPLYS
Dd	340	grgqgektsvsivr	aseeqrlreitea	eeqgghrnp	lppfrgdsr	tdtfnlllegprkia
Qy	466	NKYGAYEVKREDE	YKROLDOMDVSE	FIANITOGSMKGF	FMTNRSKVVNV	ASGEADVEMAC
Dd	460	nrtghllyaadrs	rftalndqvt	raavanlprgsml	aprylntqskl	lavllegeevqivc
Qy	536	PHL-----SGRHG	RGRGGRKHEEED	-----VHYBO	VAKLSKRE	560
Dd	460	phlgresesere	hkh--grrreede	dqgrtqrgrt	rgseseseeeq	grtvtaravrgs
Qy	561	AIIVPVGHPRVFWSS	--GNEELLFAG	INNONNHENFL	AGRENNYLOQ	IEPQAMELAF
Dd	518	afvppgrhvvleiss	qsgsnldvvc	felnaerneyw	lagr--nnvlgk	lgspsageltig
Qy	619	APRKVEELFN	SODESI--FFPGR	PHOHOOSRS	STRKQO	655
Dd	577	rparevgevfr	aqdgdegfvagp	----egs	regeqeq	610
RESULT	9					
W22149						
ID	W22149	standard; Protein: 614 AA.				
AC	W22149;					
XX						
DT	29-DEC-1997	(first entry)				
XX						
DE		Peanut allergen Ara hi.				
XX						
KW		Peanut; seed storage protein; allergen; allergy; hypersensitivity;				
KW		vaccine; anaphylactic shock; immunotherapy; therapy;				
KM		monoclonal antibody; ELISA; analysis; Ara hi.				
XX						
OS		Arachis hypogaea strain Florunner.				
XX						
EH	Key	Location/Qualifiers				
FT	Peptide	1..22				
FT	Protein	/label= Sig-peptide				
FT		23..614				
FT	Modified-site	/label= Met-Protein				
FT		521..523				
FT		/note= "N-glycosylation site"				
XX						
PN	W09724139-A1.					
PD	10-JUL-1997.					
XX						
PF	23-SEP-1996;	96WO-US15222.				
XX						
PR	04-MAR-1996;	96US-0610424.				
PR	29-DEC-1995;	95US-0009455.				
XX						
PA	(UTAR-) UNIV ARKANSAS.					
XX						
PI	Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;					

OY	93	ESGPPOOQY----	CORRKEIEEEEEEVNRODPQQYECOCORQHRETPR-----H	142
Dd	26	kssppqkktlmpcagrcldgscqgepdtkdk-----acsstctkleypdrvcydpgrh	78	
OY	143	MOTCOQR--ERR-----YEKKRKQOKRYEQOREDEEKYEERMKEDNKRDPQOR	1922	
Dd	79	tgtltisrpgpetrgtgrjgpyddrtt--qprteaggrtwgapgrereted-wrqr--	13333	
OY	193	EVEDCRRCBOEOPRQYQCOCRRCREOQRHGREGDLINQROGSSGYEEGEKOSDNPY	252	
Dd	134	--edwrlrshq-pr-----xlrpeage-----qwewlpshvreeetsnpt	175	
OY	253	yFDESLSRTFRTEBEGHISYLENFYGRSKLLRALKNRLYLLEANPAFVLPTHLADAI	312	
Dd	176	yfrstfstrfyngngrtlrvlqrdfgsrqfqnghnrivglekpvtlylphhaadnl	235	
OY	313	LIVTGRALAKMIRHDRNESYNLECGDYIRLPAGTTYLINRDNNERLHAKFLLOTISTP	372	
Dd	236	Iviqqgqalcvanagnmksfnldgnalnrlpsfisyllnnrhndqnrlrvaklsmpvtcp	2955	
OY	373	GQYKFEPFAGCONPEPYSTFSKETLEALMTQAERLRGVL-----GOOR---	417	
Dd	296	gfefdfirassrdqsylyqfslntlleaaflnaelnlrryllleenagdeeqeergqrwt	355	
OY	418	-----BCVIISASOEQLRELTRDSDSRKMHIRRGCESSRG----PYNLFNKRPlySNK	467	
Dd	356	rseennegvlvkakeheveeltkhaks-----vakkgseeagdltltpnliregeplsm	410	
OY	468	YGQAEVKEPEDYR-QLODMDVSYTIANITOGSMNGPFENRSTRKYVVYVVASGEDAVENACP	526	
Dd	411	fghtlevprdkknbpjlgdlmmlltcvelkegalmlpfniskamylvvvnkvqtnlelvay	470	
OY	527	HLSGHRGGRRGGRKHEEDEDY---HYEQVKARLSKREALIVPVGHPRVPVEVSSGNELL	5823	
Dd	471	rkeqqrgrtreedeedeedeegnrrvrytlarlkegdvflmpaabpvainasselhl	529	
OY	583	FATGTNAONNENFLACGERVNLQOIERQAMELAFAAPRKVEBELFNQSODESIFFPQPRO	642	
Dd	530	-gfignaenhnrliladkdhvldiqlekqakdiafpsgsqevexlkncqeshfvsaarpq	588	
OY	643	HOQGSRRTKOOOP	656	
Dd	589	sqsqspspeksp	602	
RESULT	14			
W62838				
ID	W62838	standard; Protein; 605 AA.		
XX	AC			
XX	W62838:			
DT	27-OCT-1998	(first entry)		
XX	XX	Glycine max antimicrobial protein.		
XX	XME			
KW		antimicrobial protein; Infestation; control.		
XX	XX			
OS		Glycine max.		
XX	PM	W09827805-A1.		
PD	02-JUL-1998.			
XX	XX			
FE	22-DEC-1997;	97WO-AU00874.		
XX	PR	20-DEC-1996;	96AU-0004275.	
XX	PA	(REFR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		
PI	Bower NI,	Goulter KC,	Green JL,	Manners JM,
XX	XX			Marcus JP.
XX	WPL:	1998-377279/32.		

PR Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
 PR useful for controlling microbial infestations of plants or mammals
 XX
 PS Claim 1; Page 63-65; 96pp; English.
 XX
 CC The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 XX
 SQ Sequence 605 AA;

Query Match	24.28;	Score 855.5;	DB 19;	Length 605;
Best Local Similarity	31.48;	Pred. No. 1.8e-68;		
Matches 194;	Conservative 143;	Mismatches 211;	Indels 69;	Gaps 15

OY	72	YDNODPTDCCOQRCRROESGPRPOOQCQRK-----KEICEEE---EYNRRDP	122
Dd	27	ywekenpkhn--kclgsnsersdyrrga-charcnllkeeeceegelprrpprpb	83
OY	123	QOOYECCSORCMIHTEFRIMQTCQRCERRRYEKERKOORRYEEOREDEKEYERMKE	182
Dd	84	eerepqpekeeedeedeprrfprrpprrpgeehnegveeqewprkeekrgkgseeded	143
OY	183	EDNKRDRPQREVIEDCRRCCEQDERPOOYCQRCRREQDGRGOGDLINORGSGSYTEE	242
Dd	144	eedeederqfiffr--prphkkearneeaded--eeqrtae-----ee	181
OY	243	GEEKO----SDNYEFEDERSLRFRTHEGISVLNFYFGSKSLRLALKNYRLVLEAN	297
Dd	182	sedeselrtthkknpflfgsnrfellknqygrltvldqfrtspqidltydltlefnsk	241
OY	298	PNAFYLPTHLDADALLLVTTGGRGALKMIIHRDNRESYMLECGDVIRIPAGTFYLLINDNN	357
Dd	242	pcllllprhadeadyllivllngtalislvinddrdsyrlgsgdalrvpsgtlyvvvpdmn	301
OY	358	ERLNIAKLQITSPGOCKEFPFPGGONPREPYLSTESKEKILEALNQARLGVL----	413
Dd	302	enllrltlalrpnkprrtesfflstceagysylqfarnlleasaydkfeelnkvlsre	361
OY	414	-----GGOR--EGVIISASOEIREFLTRDDSESRMMHIRKGESSRGPYMLFNKRPLYSN	466
Dd	362	eggqggqeqlgesavlvelskeqitalakrsksarkl----sedckflllrardplysn	417
OY	467	KYGQAYEVKRPEDYKQLDDMDYSVTIANITGSMNGPFRTNRSTVVVVVVGSGADVEMACP	526
Dd	418	k1rgkfelftpkrnpqrdldflslvmneega11prlfnskav1lvinsgdanlelv--	475
OY	527	HLSGHNG--RRGCKRHNEEDEVDYEDOKARLSKREAIVPVCHPVVFSSGMENLLFA	584
Dd	476	-----glkqgdeqqgeeplevrkyraelsesqdflvlpagyrvvv--nalsnliffa	526
OY	585	FGINAQNHNHFVLAGRENVLOQITEPOMALEAFAPRKVEBELFNSODESIFPPGPROHQ	644
Dd	527	Iginaemqnrflagsqvlnvisqrpqvqelafrgsaqavekl1knqresyfvdaqbkxk	586
OY	645	QQSSRSRTKQOQPLVSIL	661
Dd	587	eeqnkyrk--grlssll	601
<hr/>			
RESULT	15		
Y40999			
ID	Y40999	standard; protein; 605 AA.	
XX	XX		
AC	XX		
XX	Y40999;		
DT	06-DEC-1999	(first entry)	
XX			
XX	Soybean beta-conglycinin protein sequence.		
XX	Peanut; allergen; Ara H 1; IGE; immunoglobulin E; epitope; Ara h 3;		
FW	allergic reaction; soybean; beta-conglycinin		

XX Glycine max.
 OS
 XX MO9945961-A1.
 PN
 XX 16-SEP-1999.
 PD
 XX 12-MAR-1999; 99WC-US05494.
 PF
 XX 12-MAR-1998; 98US-0077763.
 PR
 XX 11-MAR-1999; 99US-0077763.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 XX
 PI Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
 PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
 XX
 DR WPI; 1999-551218/46.

PT Tertiary structure of peanut allergen Ara h 1 for protection of a host
 PT animal from allergic reaction -
 XX
 XX

PS Disclosure; Fig 33A-B; 193pp; English.

CC The invention provides a tertiary structure for the peanut allergen
 CC Ara H 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
 CC epitopes. The invention also provides an isolated recombinant peanut
 CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
 CC allergen Ara h 3. Molecules of the invention are used to protect a host
 CC animal from allergic reaction, particularly using a modified allergen
 CC which is less reactive with IgE. The invention may also be used to
 CC ensure that the allergen is not introduced into genetically modified
 CC food. The present sequence represents a soybean beta-conglycinin protein
 CC sequence.
 CC
 XX
 XX

SQ Sequence 605 AA;

Query Match 23.8%; Score 839.5; DB 20; Length 605;

Best Local Similarity 30.9%; Pred. No. 5e-67;
 Matches 192; Conservative 142; Mismatches 212; Indels 75; Gaps 15;

QY 69 WSKYNDDDPOTDCQOCRCRQOESGPROOYQORCKELCEFEERYN-----RQ 119
 DB 28 weknpxhn-----kclqscnserdsyrnqa-charcnllkveckekkgelprprp 80
 QY 120 RDPQOQYEOCQERCOHETPRHMQTCQOCERREYKERRKROKRYEQOQREDEKYEER 179
 DB 81 qhperetpqyqekedeedegprpriprrpqprgeehneqreeg---ewprkeekigekg 136
 QY 180 MKEEDNKRDPQOQREYE-DCRRRCQOEPPOOYQOCRCRQOQROHGRGDLINPQRGSG 238
 DB 137 seededededeqgerqrfprprpqhkeeneeded--eeqres----- 179
 QY 239 RYEGGEKQ-----SDNPYFDERSLSTRTEEGHISVLENFYGRSKLLRALKNRVL 293
 DB 180 --eesedseirhnknpflfsgnrfelrknygyrivrqlnqrxpsqlnrdyrlle 237
 QY 294 LEAPNFAVPTLHDADAILVTGGRCALMKIHNRNRESYNLECGDVIRIPAGTTFYLLN 353
 DB 238 fnsfpnclllpnhadayllvlnqtaislvnndrdsyrlsgdaltvpsgtlyyvn 297
 QY 354 RDNNERLHAKFLQTIISTPQOYKEFFPAGQNPPEYLSFSKEILEALNTQAERLGVL 413
 DB 298 pdnmenlrlltlaipnkpgrfseflsrtqgsyrlqfsmllleaydtkfeelinkvl 357
 QY 414 -----GOOR--EGVITASQEOJRELTRDQSESRKWHIRGGESSRGPNLNFNKR 462
 DB 358 fsreegqgqgeqrlqesviveisqelralskrakssrktl----ssedkpfnlrsrdp 413
 QY 463 LYSKKGQAEVPEDEPRLQDMDVSVFINITNIOGSMGPFNTRTKRVVVASGEADVE 522
 DB 414 lsnkikgkffeipcknpqrlididifislvdmegalllphfnskaivlivinegdanle 473

QY 523 MACPHLSGRHGC--RGGKRRHEEEDVHYEQVARKLSKREAIWVPVGHVPVSSGNEL 580
 DB 474 lv-----glkeqgqeqqgeqplevkvkryraelsqdlfviipagypvvv--natsnl 522
 QY 581 LRAFGINAONNHENFLAGKERNVLAQIEPQAMELAFAPRKEVEELFNSODESIFPPGP 640
 DB 523 nftaiginaenqgnflagsqdhnvisqipsqvgelafpsaqavkllknqrksyfvdaq 582
 QY 641 RQHQQQSSRSTKQOQPLVSTL 661
 DB 583 pkkeegnkgrk--gpissll 601

Search completed: March 1, 2001, 15:47:07
 Job time: 232 sec

